Supplementary Information

PSF decomposition of nanoscopy images via Bayesian analysis unravels distinct molecular organization of the cell membrane

Carlo Manzo,[1] Thomas S. van Zanten,[1] Suvrajit Saha,[2] Juan A. Torreno-Pina,[1] Satyajit Mayor,[2] and Maria F. Garcia-Parajo[1,3,*]

[1]ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain.

[2]National Centre for Biological Sciences (TIFR), Bellary Road, Bangalore 560065, India

[3]ICREA-Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

[*]Corresponding Author

Supplementary Methods

Detailed description of the algorithm

STED and synthetic images were analyzed by means of PSF decomposition via Bayesian analysis. The algorithm flow diagram is shown in Fig. 1c of the main text. The algorithm relies on the detection of fluorescence features and their subsequent reconstruction as a sum of PSFs whose width and intensity distribution are estimated by analyzing images of sparse markers (Supplementary Fig. S1). To ensure that the membrane environment does not alter the intensity distribution of the markers, we compared the distribution obtained from markers immobilized on a glass coverslip with that obtained by imaging the labeled mutated DC-SIGN receptor (ΔRep), mostly organized as monomers on CHO cell membranes[1]. The similarity between the two distributions validates our calibration for marker intensity (Supplementary Fig. S1c).

A schematic view of the working algorithm is presented in Supplementary Fig. S2.

The algorithm is applied on the raw images without filtering or preprocessing. The first step consists in the SEARCH routine used to determine the image area having the highest

probability of containing features that can be associated to marker fluorescence emission (upper row on Supplementary Fig. S2). The routine calculates the likelihood that the pixel intensity values $I_{i,j}$ in a square box of size b_w have a Gaussian (noise-like) distribution with background offset m and variance s^2:

$$L_{noise} = \prod_{i,j=1}^{b_w} \exp\left\{ -\frac{\left(I_{i,j} - \hat{m}\right)^2}{2\hat{s}^2} \right\}, \qquad \text{(eqn. S1)}$$

where the hat symbols indicate maximum likelihood estimators (MLEs). MLEs are obtained by partial derivation of the likelihood function, giving:

$$\hat{m} = \frac{1}{b_w^2} \sum_{i,j=1}^{b_w} I_{i,j}, \qquad \text{(eqn. S2)}$$

$$\hat{s}^2 = \frac{1}{b_w^2} \sum_{i,j=1}^{b_w} \left(I_{i,j} - \hat{m}\right)^2. \qquad \text{(eqn. S3)}$$

The calculation is performed over the whole image by sliding the box pixel by pixel. A box size of length b_w = 4×FWHM+1 was found to correctly estimate the presence of emitters and limit false positive identifications. Boxes partly lying outside the image area are excluded. In order to minimize computational time, the calculation of L_{noise} and MLEs is performed via Fourier transform. In this way, we obtain a likelihood map of the image (Supplementary Fig. S2). The pixel position showing the maximum negative likelihood - L_{noise} (corresponding to the maximum probability of having features that do not correspond to noise) is singled out and the box centered at this pixel position is passed to the next algorithm step, consisting in the box-*BIC* evaluation (Supplementary Fig. S2, middle row). In this step, the intensity map of the box region is attempted to be reconstructed as the sum of *n PSF*s. Each *PSF* corresponds to a two-dimensional Gaussian function:

$$PSF(x_0, y_0, I_0) = I_0 \times e^{\frac{-(x-x_0)^2-(y-y_0)^2}{2\sigma^2}}, \quad \text{(eqn. S4)}$$

where x_0 and y_0 are the peak location coordinates, I_0 is the peak intensity and σ is the radius at which the *PSF* decays at ~60% of its maximum. For a given image, σ is kept fixed at the value

$$\sigma = \frac{FWHM}{\sqrt{8\log(2)}}, \quad \text{(eqn. S5)}$$

where the *FWHM* is independently measured from images of sparse markers adsorbed on a coverslip and sets the resolution of the STED image (Supplementary Fig. S1a-b). The peak coordinates and intensity are used as free parameters and the fitting is performed via minimization of the Bayesian Information Criterion:

$$BIC(n) = -2\log\left(\prod_{i,j=1}^{b_w} L_{rec}(I_{i,j}) \prod_{k=1}^{n} L_{int}(I_k)\right) + 3n. \quad \text{(eqn. S6)}$$

In the last equation, the first product in the logarithm represents the likelihood of the model in reconstructing the image and is given by:

$$L_{rec} = \exp \frac{-\left(\sum_{k=1}^{n} PSF(x_k, y_k, I_k) - I_{i,j} - M\right)^2}{2S^2}, \quad \text{(eqn. S7)}$$

with M and S^2 being the local estimation for noise background offset and variance.

The second term in the logarithm describes the likelihood for the *PSF* peak intensities to belong to the marker intensity distribution:

$$L_{int} = g_I(I_k), \quad \text{(eqn. S8)}$$

where g_I is the probability distribution of marker peak intensities experimentally determined from the image of sparse antibodies (Supplementary Fig. S1a). To allow

calculation of g_I for every value of I_k, the peak intensities histogram obtained from isolated markers was interpolated by means of a Gaussian kernels density distribution (Supplementary Fig. S1c).

The third term in Eqn. S6 introduces a penalty for the addition of further *PSF*s, thus preventing overestimation of the particle number. The number of *PSF*s used to model the intensity distribution within the box is sequentially increased and the corresponding *BIC* calculated until a stable global minimum is obtained, with $n_{\max} = \arg_n\{\min(BIC)\} + 2$.

The *PSF*s corresponding to the *BIC* minimum are considered as those providing the most faithful reconstruction of the intensity map of the box and are stored for the subsequent analysis. Although during the fitting routine the *PSF*s are allowed to have center positions lying even outside the box, only the coordinates of the *PSF*s located at a distance from the box edges larger than $b_w/6$ are passed to the next algorithm block, the REC/SUB routine. If no *PSF*s are found within this area, the value of L_{noise} of the box center pixel is set to 0 to prevent the algorithm to return to evaluate the same area.

In the REC/SUB routine, the *PSF*s found by box-*BIC* are subtracted from the original image (SUB) and in parallel added to a null matrix to obtain the partial reconstructed image (REC, Supplementary Fig. S2, lower row). Then, the SUB image becomes the new entry of the algorithm in the next iteration as schematically indicated by the red dashed arrows in Supplementary Fig. S2. At each iteration, newly found *PSF*s are progressively added to the REC image and subtracted from the SUB one. After every REC/SUB evaluation, the global likelihood of the subtracted image intensity is evaluated and the algorithm is stopped when further subtractions of PSFs does not cause an increase in the likelihood value. Although this choice allows fast calculation, it tends to slightly overestimate the number of emitters and to produce a higher false positive recognition rate. Therefore, the routine is finally refined by the application of cumulative *BIC* analysis on the entire reconstructed image. In this finishing step, all the individual localizations

previously retrieved are sorted in descending order of their individual likelihood. The *BIC* is then cumulatively calculated over the whole image. The localizations producing a decrease of the overall *BIC* up to a global minimum are retained, whereas those producing a *BIC* increase are excluded (Supplementary Fig. S2, most right-side upper panel).

In order to perform the *BIC* calculation in eqns. S6 and S7, the algorithm requires the estimation of mean and variance of the underlying noise, M and S^2. Although these parameters can be evaluated from image areas free of fluorescent markers, in our analysis we opted for an automatic evaluation, based on the application of the described method without taking into account intensity constraints on the PSF. Such a choice allows to build up a background noise map, taking into account local changes in background noise due to cell autofluorescence.

The algorithm was written in Matlab (The MathWorks, Inc.) and all the analysis were performed on a single core i7 processor (3.40 GHz) computer.

Quantification of receptors spatial organization

Receptor spatial organization was quantified by calculating the *nnd*[2] and the pair-correlation function[3,4], from the localization coordinates obtained via application of the algorithm to the STED images and taking into account the related localization accuracy.

Calculation of the *nnd* was performed by generating a distance matrix $r_{i,j}$ from the localization coordinate list (x_i,y_i), with elements given by $r_{i,j} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}$. The minimum of each column for $i<j$ represents the distance between the j-th localization and its closest neighbor i^* and was used to build the *nnd*. The distribution was calculated by taking into account that, due to the finite localization accuracy, the distance between localized receptors has an error $\Delta r_{i,j}$ given by $\Delta r_{i,j} = \sqrt{\Delta r_i^2 + \Delta r_j^2}$, where Δr_i is the

accuracy of the *i*-th localization, thus providing $nnd(r) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{\sqrt{2\pi}\Delta r_{i^*,j}} e^{\frac{-(r-r_{i^*,j})^2}{2\Delta r_{i^*,j}^2}}$, where *N* represents the number of localizations.

The *nnd*s obtained for the investigated receptors were compared with simulations of randomly distributed particles (uniform spatial distribution) at the same surface density.

The accuracy at which each localization is determined depends on several factors, such as PSF width, number of photons, pixel size and background level[5] as well as the local density of emitters in the neighboring region and their distances[6]. In order to take into account the variation in molecular density present in our experiments, we estimated the localization accuracies Δr_i as 3 times the Cramer-Rao lower bound for an isolated Gaussian having same PSF width, number of photons, pixel size and background level[5]. The average values $\langle \Delta r_i \rangle$ obtained for our experiments are higher than the maximum values obtained from simulations (Figure 3 of the main text), thus our choice represents a conservative estimate of the localization accuracy. For the simulations, we assumed a constant localization accuracy of 25 nm, corresponding to the average localization accuracy obtained for our experiments. Enhanced proximity probability was calculated as the integral of the difference of the experimental and simulated *nnd*s between *r*=0 and the first cross point between the two distributions.

The pair-correlation function *g*(*r*) was calculated similarly as previously detailed[3,4] on reconstructed images obtained by convoluting the particle localizations with 2-d Gaussian functions having width equal to the localization accuracies Δr_i. Each curve represents the average of at least 30 pair correlations obtained over 4×4 μm^2 regions of interest extracted from at least 3 cells per condition.

Fitting of the *g*(*r*) curves was performed according to:

$$g(r) = \frac{1}{4\pi\langle \Delta r_i \rangle^2 \rho}\exp\left(-\frac{r^2}{4\langle \Delta r_i \rangle^2}\right) + g_{rec} \otimes g_{\langle \Delta r_i \rangle},$$

where the first term accounts for the average particle localization accuracy $\langle \Delta r_i \rangle$ and with ρ indicating the molecular density. The second term represents the convolution of the receptor correlation with a Gaussian function having width equal to the average localization accuracy. The data obtained for the FR-GPI and the FRTM-Ez-AFBD receptor were fitted by means of an exponential function $g_{rec} = A\exp\left(-r/\xi\right)$, with ξ being the correlation radius. For the FRTM-Ez-AFBD* case, the correlation function also included a cosine term, $g_{rec} = A\cos\left(\pi r/2r_0\right)\exp\left(-r/\xi\right)$, with the parameter r_0 estimating the average cluster separation. The fitting was performed on the parameters A, ξ and r_0, whereas $\langle \Delta r_i \rangle$ and ρ were fixed at the values determined by the PSF decomposition ($\langle \Delta r_i \rangle$= 22, 18 and 27 nm and ρ = 43, 18 and 11 μm^{-2} for FR-GPI, FRTM-Ez-AFBD and FRTM-Ez-AFBD*, respectively). The cluster radius r_{clust} was calculated as the value at which g_{rec} decays at $1/e$. From these quantities, the average number of particle per cluster N_{clust} was calculated as described in Sengupta et al[4]. The errors on the fitting parameters were calculated as the 95.4% confidence interval[7]. Errors on the derived quantities were obtained through statistical error propagation. All the analyses were performed using custom routines written in Matlab (The MathWorks, Inc.).

Comparison with Compressed Sensing method

In order to compare the performance of our method to other algorithms, we have analyzed a set of our simulated data with the previous reported Compressed Sensing method for STORM[8] using the Matlab code provided by its authors. We used a patch size of 7x7 pixels, ε=1.5 and a 12x12 subdivision (corresponding to 2.1 nm per pixels). In order to allow comparison with our method, compressed sensing results were converted into lists of molecular positions[8].

Supplementary Figure 3 compares the rate of molecule identification (recall fraction, R_f) of the two methods as a function of the molecular density for several STED

parameters. The larger use of prior knowledge (probability distribution of marker intensity and instrumental PSF) in our Bayesian method provides superior performance in the investigated range of parameters.

The difference in instrumental settings (pixel size, detectors, fluorophores) between STORM and STED imaging prevent a direct comparison of our results with those previously published[8].

1 Manzo, C. *et al.* The neck region of the C-type lectin DC-SIGN regulates its surface spatiotemporal organization and virus-binding capacity on antigen-presenting cells. *J. Biol. Chem.* **287**, 38946-38955, (2012).
2 van Zanten, T. S. *et al.* Hotspots of GPI-anchored proteins and integrin nanoclusters function as nucleation sites for cell adhesion. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 18557-18562, (2009).
3 Veatch, S. L. *et al.* Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PLoS ONE* **7**, (2012).
4 Sengupta, P. *et al.* Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Methods* **8**, 969-975, (2011).
5 Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775-2783, (2002).
6 Huang, F., Schwartz, S. L., Byars, J. M. & Lidke, K. A. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. *Biomed. Opt. Express* **2**, 1377-1393, (2011).
7 Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical recipes in C (2nd ed.): the art of scientific computing*. (Cambridge University Press, 1992).
8 Zhu, L., Zhang, W., Elnatan, D. & Huang, B. Faster STORM using compressed sensing. *Nat. Methods* **9**, 721-723, (2012).

SUPPLEMENTARY FIGURES

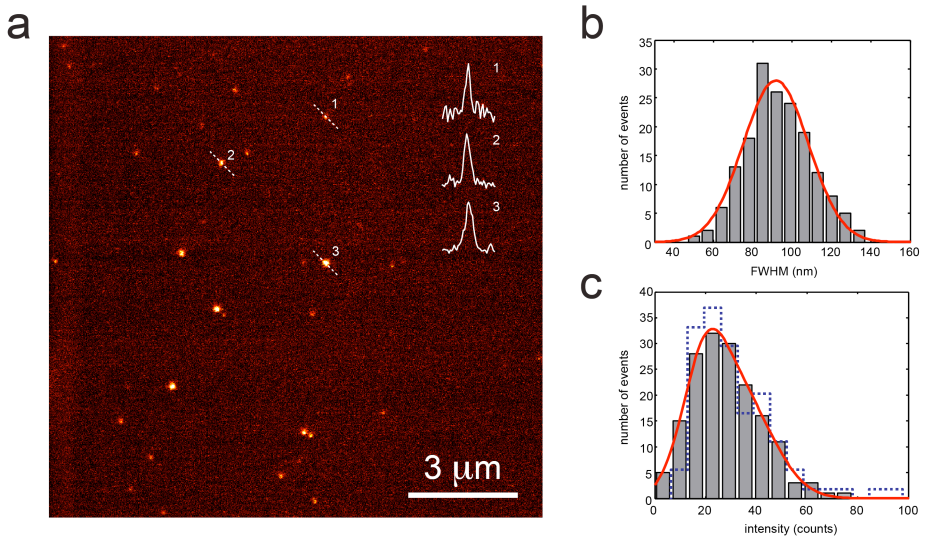

Figure 1. PSF and intensity distribution of single fluorescent markers. (a) STED image of single fluorescent markers adsorbed on a glass coverslip. Normalized line profiles of fluorescent markers show the PSF shape and width. (b) Distribution of the FWHM of several fluorescent markers allow determination the instrumental resolution of the STED microscope, providing 93±15 nm (s.d.). Red line represents the Gaussian fit of the distribution. (c) Distribution of peak intensity of single fluorescent markers immobilized on a glass coverslip (bars) and for labeled DC-SIGN receptor mutants (ΔRep)[1] on CHO cell membranes (dotted line). The continuous distribution (red line) was estimated by a Gaussian kernel function.

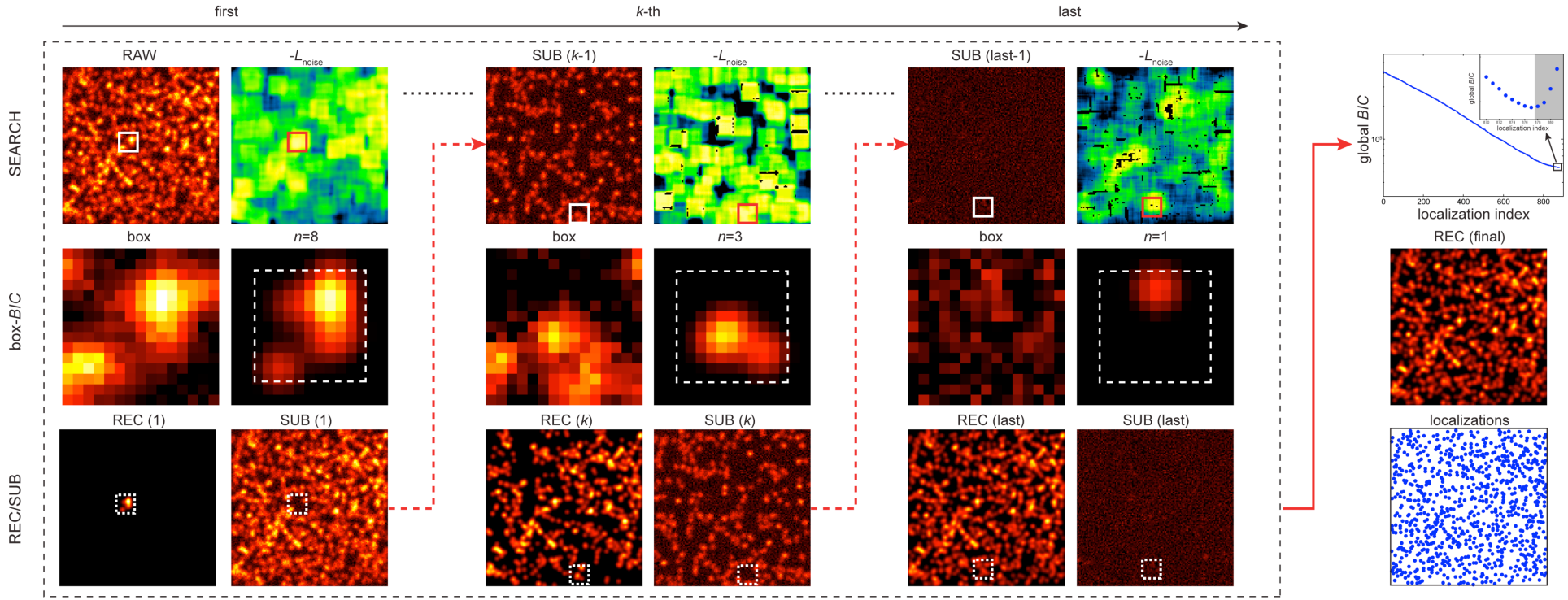

Figure 2. Schematic of the algorithm working flow.

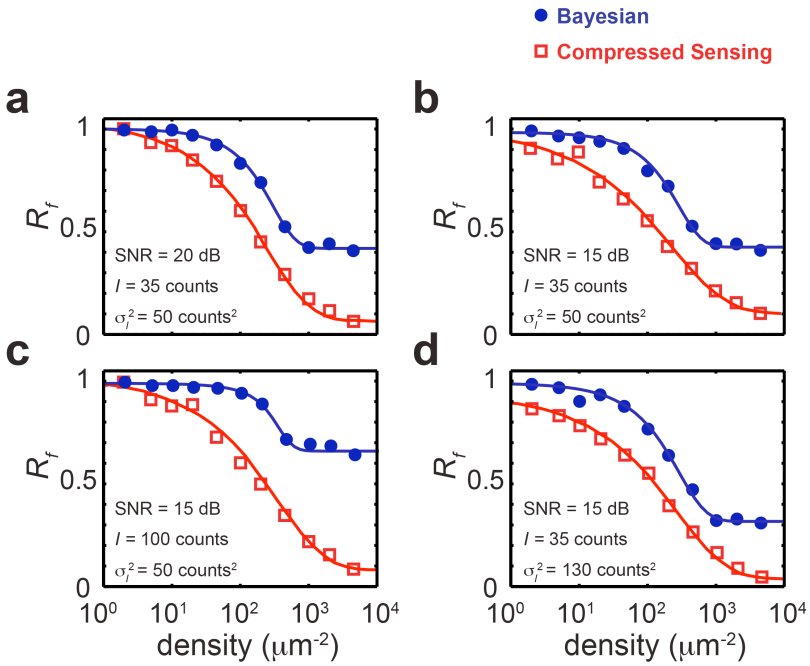

Figure 3. Comparison of algorithm performance with the compressed sensing method of Zhu *et al*[8]. The recall fraction R_f as a function of molecular density for simulated images with FWHM=90 nm at varying imaging parameters. For the compress sensing analysis we used a patch size of 7x7 pixels, ε=1.5 and a 12x12 subdivision (corresponding to 2.1 nm per pixels). In order to allow comparison with our method, compressed sensing results were converted into lists of molecular positions[8].